BECCA: A BICA for arbitrary robots in unknown worlds

Brandon Rohrer* **J. Dan Morrow, Fred Rothganger, Patrick Xavier**
Sandia National Laboratories
Albuquerque, New Mexico, USA

Natural world interaction This work is part of an extended effort to improve AI systems' performance on *natural world interaction* tasks, that is, tasks that typically occur in the physical world, whether in nature or in man-made environments. These tasks are widely variable and require interaction with environments that are, in general, stochastic, non-stationary, non-Markovian, partially observable, and highly non-linear. Natural world interaction can be represented as a general discrete time reinforcement learning problem in which an agent makes periodic observations, selects actions, and receives a reward signal. Although it is difficult to make performance guarantees for any approach to the problem of natural world interaction formulated in this way, partial solutions to this problem have been put forward in the form of biological organisms. Behavioral and neurological studies yield a rich set of insights and clues as to how this problem might be addressed in machines.

BECCA One biologically-inspired approach is the Brain-Emulating Cognition and Control Architecture (BECCA, see Figure 1). BECCA issues action commands, takes in observations (which includes a copy of the actions), and receives rewards based on its performance.

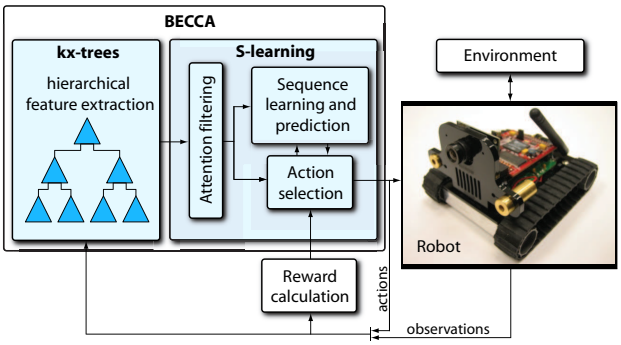

Figure 1: Block diagram of BECCA applied to an arbitrary robotic task.

BECCA performs feature extraction on its observations using a novel unsupervised learning method, **kx-trees**. It works by online divisive partitioning of the state space into regions, based on the local frequency of observations. Each region that contains a sufficiently large number of observations is a feature. A new state space can then be created, using the features as inputs, and the partitioning and feature extraction process can be repeated at a higher level of abstraction. This process can be repeated as many times as the data supports and as computational resources allow.

Reward-driven, model-based action selection is performed by **S-learning**. Each observation may match a number of features. An attention filter selects among these based on salience, unexpectedness, and prior association with reward, and passes one feature each time step to a sequence learning process. Over time, repeated sequences are weighted more heavily and unrepeated sequences are forgotten. The resulting sequence library constitutes a model of the robot and the environment. Given an initial state, the library also shows which states can be reached by which sequence of actions. Actions are selected based on the expected reward of the predicted states.

BECCA learns without making use of prior knowledge. S-learning and kx-trees avoid making common assumptions about their data. They do not assume that the data values vary proportionally, smoothly, continuously, or even that they are ordered fields. As a result, BECCA can use any type of data, continuous, discrete, or symbolic, including vision, audio, contact, distance, text, or processor function calls.

Mobile robot application BECCA was implemented on a Surveyor SRV-1 robot (Surveyor Corp, CA, USA), with a 24-dimensional continuous visual state space and 4 possible actions (forward, back, right, and left) to choose from. It was tasked to find alternately high- and low-contrast visual scenes.It learned how to do this reasonably well within one hour and within 12 hours had settled on a nearly optimal solution. BECCA did this, despite having no knowledge of the structure of the robot or its environment.

Acknowledgements This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy under contract DE-AC04-94AL85000.

*Corresponding author. Email: brrohre@sandia.gov, Internet: www.sandia.gov/∼brrohre/